

Méthodes numériques en finance

Nicolas Baradel

22 juin 2020

Table des matières

Introduction	1
1 Mesure de l'erreur de Monte Carlo	3
2 Simulation de variables aléatoires	4
3 Méthodes de réduction de variance	6
3.1 Variable antithétique	6
3.2 Variable de contrôle	7
3.3 Fonction d'importance ou échantillonnage préférentiel	8
3.4 Méthode de stratification	9
4 Simulation de processus stochastiques	12
4.1 Simulation d'un mouvement brownien	12
4.2 Simulation d'une équation différentielle stochastique	13
4.2.1 Simulation exacte de processus d'Itô	13
4.2.2 Schéma d'Euler	15
4.2.3 Schéma de Milstein	16
4.3 Application à l'évaluation de produits dérivés	16
5 Méthodes des différences finies	17
5.1 Formule de Keynman-Kac	17
5.2 Schéma explicite	17
5.3 Schéma implicite	19
5.4 θ -schéma	20
A Simulations - Méthode du rejet	21
B Formule de Black Scholes	23
C Exercice 11 - Prix des Zéros Coupon dans le modèle de Vašíček	23

Introduction

L'objectif est de calculer ou d'estimer des quantités de la forme :

$$\theta := \mathbb{E}[f(X)] = \int_{\mathbb{R}^d} f(x) d\mathbb{P}_X(x),$$

où

- X est une variable aléatoire à valeurs dans \mathbb{R}^d ,
- f est une fonction mesurable de \mathbb{R}^d dans \mathbb{R} ,
- \mathbb{P}_X est la loi de la variable aléatoire X ,
- $\mathbb{E}[|f(X)|] < +\infty$.

La méthode la plus simple est la **méthode de Monte Carlo**. Cette méthode consiste à simuler une suite (X_1, \dots, X_n) dont les composantes ne sont pas obligatoirement indépendantes, mais telles qu'une loi des grands nombres s'applique, c'est à dire que :

$$\hat{\theta}_n := \frac{1}{n} \sum_{k=1}^n f(X_k) \xrightarrow[n \rightarrow +\infty]{p.s.} \theta.$$

Une telle convergence s'obtient grâce au théorème de la loi forte des grands nombres :

Théorème 1 (Loi forte des grands nombres). *Soit $(X_k)_{1 \leq k \leq n}$ une suite i.i.d. de variables aléatoires réelles telle que $\mathbb{E}(|X_1|) < +\infty$. Alors*

$$\frac{1}{n} \sum_{k=1}^n X_k \xrightarrow[n \rightarrow +\infty]{p.s.} \mathbb{E}(X_1).$$

On y associera un théorème central limite qui permet d'avoir une mesure de **l'erreur d'estimation**.

En finance, nous sommes souvent face à une variable aléatoire X qui est une semi-martingale d'Itô définie par une équation de la forme :

$$X_s^{t,x} = x + \int_t^s \mu(s, X_s^{t,x}) ds + \int_t^s \sigma(s, X_s^{t,x}) dW_s,$$

où :

- $x \in \mathbb{R}^d$,
- $\mu : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$,
- $\sigma : [0, T] \times \mathbb{R}^d \rightarrow S_d^+(\mathbb{R})$,
- W est un mouvement brownien,

et de telle sorte que l'équation ci-dessus ait une unique solution.

Ce cas ouvre la voie à une seconde méthode d'estimation de θ . Dans ce cas, la formule de Feynman Kac permet de relier le calcul d'intégral à une équation aux dérivées partielles qu'il est possible de résoudre numériquement.

Le cours se découpe de la façon suivante. Dans un premier temps nous mettrons en place les éléments qui permettent d'en déduire une mesure de l'erreur d'estimation par Monte Carlo (construction d'intervalles de confiance). Ensuite nous verrons des méthodes de simulation de variables aléatoires. Nous introduirons le concept de réduction de variance et nous verrons différents contextes. Nous simulerons ensuite des processus stockastiques afin d'appliquer les précédentes méthode au calcul du prix d'actifs dérivés. Enfin, nous utiliserons les schéma numériques sur les équations aux dérivées partielles afin de calculer le prix d'actifs dérivés.

Les applications se feront sur Python.

1 Mesure de l'erreur de Monte Carlo

Il est essentiel d'être capable de mesurer l'erreur commise par une méthode d'estimation. Pour les méthodes de Monte Carlo, celle-ci repose souvent sur le théorème central limite.

Théorème 2 (Théorème central limite). *Soit $(X_k)_{1 \leq k \leq n}$ une suite i.i.d. réelle telle que $\mathbb{E}(X_1^2) < +\infty$. Alors*

$$\sqrt{n} \left(\frac{1}{n} \sum_{k=1}^n X_k - \mathbb{E}(X_1) \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \text{Var}(X_1)).$$

Grâce à ce théorème, nous pouvons définir un intervalle de confiance asymptotique

Proposition 1. *Soit $\alpha \in]0, 1[$. Soit $\Phi^{-1}(\alpha)$ le quantile de niveau α de la loi normale. Soit :*

$$\begin{aligned} IC_{1-\alpha}^n[\mathbb{E}(X_1)] &= \left[\bar{X}_n \pm \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \frac{\sigma}{\sqrt{n}} \right] \\ &= \left[\bar{X}_n - \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \frac{\sigma}{\sqrt{n}}, \bar{X}_n + \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \frac{\sigma}{\sqrt{n}} \right], \end{aligned}$$

où $\bar{X}_n := \frac{1}{n} \sum_{k=1}^n X_k$ et $\sigma^2 := \text{Var}(X_1)$. Il s'agit d'un intervalle de confiance asymptotique sur $\mathbb{E}(X_1)$ de niveau $1 - \alpha$ c'est à dire :

$$\mathbb{P}(\mathbb{E}(X_1) \in IC_{1-\alpha}^n[\mathbb{E}(X_1)]) \xrightarrow[n \rightarrow +\infty]{} 1 - \alpha.$$

Le paramètre σ est généralement inconnu. On peut l'estimer avec :

$$\hat{\sigma} := \sqrt{\frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X}_n)^2}.$$

Remarque 1. *Dans Python, la fonction `std` du package `numpy` divise par n tandis que la fonction `sd` de R divise par $n - 1$.*

L'intervalle de confiance précédent reste valide asymptotiquement par application du théorème de Slutsky car $\hat{\sigma}$ est convergent.

Il est également possible de construire un intervalle de confiance sur $g(\bar{X}_n)$ avec g suffisamment régulière.

Théorème 3 (Méthode delta). *Soit $(X_k)_{1 \leq k \leq n}$ une suite de variables aléatoires réelles telle que*

$$\sqrt{n} (\bar{X}_n - \theta) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2),$$

et $g : \mathbb{R} \rightarrow \mathbb{R}$ une fonction C^1 en θ et $g'(\theta) \neq 0$. Alors,

$$\sqrt{n} (g(\bar{X}_n) - g(\theta)) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2 g'(\theta)^2).$$

Exemple 1. *Soit $(X_k)_{1 \leq k \leq n}$ une suite i.i.d. de loi de $\mathcal{P}(\lambda)$. On souhaite estimer $\mathbb{E}(X_1) = \lambda$.*

Alors $IC_{1-\alpha}^n[\lambda] = \left[\left(\sqrt{\bar{X}_n} \pm \frac{\Phi^{-1}(1-\frac{\alpha}{2})}{2\sqrt{n}} \right)^2 \right]$ est un intervalle de confiance asymptotique de niveau $1 - \alpha$. On a ici stabilisé la variance : elle ne dépend plus de la quantité inconnue à estimer.

Démonstration. Par le théorème central limite :

$$\sqrt{n} (\bar{X}_n - \lambda) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \lambda).$$

On pose $g : x \mapsto \sqrt{x}$. On a $g'(x) = \frac{1}{2\sqrt{x}}$. En appliquant la méthode Delta, on en déduit :

$$\sqrt{n} \left(\sqrt{\bar{X}_n} - \sqrt{\lambda} \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \frac{1}{4} \right).$$

Il vient :

$$IC_{1-\alpha}^n[\sqrt{\lambda}] = \left[\sqrt{\bar{X}_n} \pm \frac{\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)}{2\sqrt{n}} \right],$$

puis, comme g est strictement croissante sur \mathbb{R}_+ ,

$$IC_{1-\alpha}^n[\lambda] = \left[\left(\sqrt{\bar{X}_n} \pm \frac{\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)}{2\sqrt{n}} \right)^2 \right] = \left[\left(\sqrt{\bar{X}_n} - \frac{\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)}{2\sqrt{n}} \right)^2, \left(\sqrt{\bar{X}_n} + \frac{\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)}{2\sqrt{n}} \right)^2 \right].$$

□

Exercice 1. Soit $(X_k)_{1 \leq k \leq n}$ une suite i.i.d. de loi $\mathcal{P}(\lambda)$ avec $\lambda > 0$. Comme dans l'exemple précédent, on souhaite estimer $\mathbb{E}(X_1) = \lambda$.

Montrez que $IC_{1-\alpha}^n[\lambda] = \left[\exp \left(\log(\bar{X}_n) \pm \frac{\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)}{\sqrt{n\bar{X}_n}} \right) \right]$ est un intervalle de confiance asymptotique de niveau $1 - \alpha$.

2 Simulation de variables aléatoires

Il existe diverses méthodes de simulations de variables aléatoires réelles. L'une d'entre elles est universelle lorsqu'on connaît l'inverse de la fonction de répartition de la loi qu'on souhaite simuler. Définissons d'abord la fonction inverse.

Définition 1. Si on note F_X la fonction de répartition d'une variable aléatoire X , l'inverse généralisée de F_X , également appelée fonction quantile de la loi de X , est définie par

$$F_X^{\leftarrow}(\alpha) := \inf \{ x \in \mathbb{R} : F_X(x) \geq \alpha \}, \quad \alpha \in]0, 1[.$$

Lorsque F_X est inversible, c'est l'inverse habituel, c'est à dire que $F_X^{\leftarrow} = F_X^{-1}$.

On a le résultat fondamental suivant.

Théorème 4. Soit $U \sim \mathcal{U}([0, 1])$. Alors

$$F_X^{\leftarrow}(U) \stackrel{\mathcal{L}}{=} X.$$

Démonstration. Dans le cas où F_X est inversible, on a

$$F_{F_X^{-1}(U)}(x) = \mathbb{P}(F_X^{-1}(U) \leq x) = \mathbb{P}(U \leq F_X(x)) = F_X(x).$$

Les deux variables aléatoires ont la même fonction de répartition, elles ont donc la même loi. Pour le cas général, il faut montrer au préalable que

$$F_X^{\leftarrow}(\alpha) \leq x \iff \alpha \leq F_X(x),$$

puis la démonstration est identique. □

Exemple 2 (Loi de Bernoulli). Soit $X \sim \mathcal{B}(p)$ avec $p \in]0, 1[$. On a $F_X(x) = (1 - p)\mathbf{1}_{[0,1[}(x) + \mathbf{1}_{[1,+\infty[}(x)$. On en déduit $F_X^{\leftarrow}(U) = \mathbf{1}_{[1-p,1]}(U) \stackrel{\mathcal{L}}{=} X$.

De l'exemple précédent, on déduit encore plus simplement que $\mathbf{1}_{[0,p]}(U) \stackrel{\mathcal{L}}{=} X$ car $U \stackrel{\mathcal{L}}{=} 1 - U$. De manière générale, pour simuler une loi discrète, il suffit de découper U dans des intervalles dont la taille est la probabilité de chaque évènement, et d'y associer l'évènement correspondant (ici 0 avec probabilité $1 - p$ et 1 avec probabilité p).

Exemple 3 (Loi exponentielle). Soit $X \sim \mathcal{E}(\lambda)$ avec $\lambda > 0$. On a $F_X(x) = (1 - e^{-\lambda x})\mathbf{1}_{\{x \geq 0\}}$. On en déduit $F_X^{\leftarrow}(U) = -\frac{1}{\lambda} \log(1 - U) \stackrel{\mathcal{L}}{=} X$.

Dans le cas continue, il suffit de calculer l'inverse classique et de l'appliquer. On peut également utiliser le fait que $U \stackrel{\mathcal{L}}{=} 1 - U$ et en déduire plus simplement que $-\frac{1}{\lambda} \log(U) \stackrel{\mathcal{L}}{=} X$.

Remarque 2. La fonction de répartition n'est pas toujours explicite. C'est le cas de la loi normale : celle-ci est approximée dans tous les langages de calcul scientifique (*norm.cdf* pour Python via le package *scipy.stats* et *pnorm* dans R, et respectivement *norm.ppf* et *qnorm* pour les inverses).

Exercice 2. Estimer via Monte Carlo les quantités suivantes :

- $\mathbb{E}[\sin(U)]$ où $U \sim \mathcal{U}([0, \pi])$
- $\mathbb{E}[\log(X)]$ où $X \sim \mathcal{E}(1)$
- $\mathbb{E}[(e^{-\frac{\sigma^2}{2} + \sigma Z} - 1)^+]$ où $Z \sim \mathcal{N}(0, 1)$ et $\sigma = 0.2$

On n'utilisera que des simulations de loi uniforme pour cet exercice. On utilisera $n = 10^6$ et on donnera un intervalle de confiance pour chacune des estimations.

Pour simuler une loi $\mathcal{U}([0, 1])$, dans Python on utilisera la fonction `random.random(n)` du package `numpy`, dans R on utilisera la fonction `runif(n)` où n est le nombre de simulations.

On pourra afficher les histogrammes de U , X et Z .

Exercice 3. Soit $(X, Y) \sim \mathcal{U}([-1, 1]^2)$.

- Montrez que $\mathbb{P}(X^2 + Y^2 \leq 1) = \frac{\pi}{4}$.
- En déduire un algorithme d'estimation de π par méthode de Monte Carlo et le mettre en place en Python.

Proposition 2 (Méthode de Box-Muller). Soit $(U, V) \sim \mathcal{U}([0, 1]^2)$. Soient

$$X := \sqrt{-2 \log(U)} \cos(2\pi V), \quad Y := \sqrt{-2 \log(U)} \sin(2\pi V)$$

Alors X et Y sont deux variables aléatoires indépendantes de loi $\mathcal{N}(0, 1)$.

Démonstration. On utilise la caractérisation de la loi suivante : deux variables aléatoires A et B à valeurs dans \mathbb{R}^d ont la même loi si et seulement si, pour toute fonction continue bornée $f : \mathbb{R}^d \mapsto \mathbb{R}$,

$$\mathbb{E}[f(A)] = \mathbb{E}[f(B)].$$

Soit $(W, Z) \sim \mathcal{N}(0, 1)^{\otimes 2}$. Alors, avec le changement de variable $(w, z) = (r \cos(\theta), r \sin(\theta))$,

$$\begin{aligned} \mathbb{E}[f(W, Z)] &= \int_{\mathbb{R}} \int_{\mathbb{R}} f(w, z) \frac{1}{2\pi} e^{-\frac{w^2+z^2}{2}} dw dz \\ &= \int_0^{2\pi} \frac{1}{2\pi} \int_0^{+\infty} f(r \cos(\theta), r \sin(\theta)) r e^{-\frac{r^2}{2}} dr d\theta \\ &= \int_0^{2\pi} \frac{1}{2\pi} \int_0^{+\infty} f(\sqrt{s} \cos(\theta), \sqrt{s} \sin(\theta)) \frac{1}{2} e^{-\frac{s}{2}} ds d\theta. \end{aligned}$$

Par ailleurs, par l'Exemple 3, $S := -2 \log(U) \sim \mathcal{E}(2)$ et $\Theta := 2\pi V \sim \mathcal{U}([0, 2\pi])$,

$$\begin{aligned} \mathbb{E}[f(X, Y)] &= \mathbb{E}\left[f(\sqrt{S} \cos(\Theta), \sqrt{S} \sin(\Theta))\right] \\ &= \int_0^{2\pi} \frac{1}{2\pi} \int_0^{+\infty} f(\sqrt{s} \cos(\theta), \sqrt{s} \sin(\theta)) \frac{1}{2} e^{-\frac{s}{2}} ds d\theta. \end{aligned}$$

□

Exercice 4. Implémentez la méthode de Box-Muller dans Python. On testera la normalité avec un test de Kolmogorov-Smirnov grâce à la fonction `kstest` du package `scipy.stats` (fonction `ks.test` dans R).

3 Méthodes de réduction de variance

Une fois la méthode de simulation choisie, on estime

$$\theta := \mathbb{E}[f(X)] = \int_{\mathbb{R}^d} f(x) d\mathbb{P}_X(x).$$

Dans le cas standard, on simule un échantillon $(X_1, \dots, X_n) \stackrel{i.i.d.}{\sim} \mathbb{P}_X$ puis on estime θ par :

$$\hat{\theta}_n := \frac{1}{n} \sum_{k=1}^n f(X_k).$$

L'estimateur est sans biais et sa précision est directement liée à sa variance qui est

$$\text{Var}(\hat{\theta}_n) = \frac{\sigma_{f(X)}^2}{n}.$$

L'objectif des méthodes de réduction de variance est de réduire $\sigma_{f(X)}^2$.

3.1 Variable antithétique

Soit X une variable aléatoire telle que $X \stackrel{\mathcal{L}}{=} a - X$ (symétrie en $a \in \mathbb{R}$). On pose l'estimateur :

$$\hat{\theta}_n^A := \frac{1}{n} \sum_{k=1}^n \frac{f(X_k) + f(a - X_k)}{2}.$$

Cet estimateur demeure sans biais et on a

$$\text{Var}(\hat{\theta}_n^A) = \frac{\sigma_{f(X)}^2 + \text{Cov}(f(X), f(a - X))}{2n}.$$

Lemme 1. L'estimateur antithétique $\hat{\theta}_n^A$ est meilleur que l'estimateur $\hat{\theta}_n$.

Démonstration. L'inégalité de Cauchy-Schwarz implique que $\text{Cov}(f(X), f(a - X)) \leq \sigma_{f(X)}^2$. Ainsi,

$$\text{Var}(\hat{\theta}_n^A) \leq \frac{\sigma_{f(X)}^2}{n} = \text{Var}(\hat{\theta}_n).$$

□

Il est à noter que l'ajout de la version antithétique implique le calcul de f à nouveau. Il est intéressant de pouvoir quantifier l'amélioration, car cette dernière n'est pas réelle, on est moins efficace.

Lemme 2. *Soit f une fonction monotone. Dans ce cas*

$$\text{Var}(\hat{\theta}_n^A) \leq \frac{\sigma_{f(X)}^2}{2n}.$$

Démonstration. Il faut montrer que $\text{Cov}(f(X), f(a - X)) \leq 0$, ce qui sera admis. □

Remarque 3. *Dans les hypothèses du lemme 2, l'estimateur antithétique est en pratique toujours plus efficace. En effet, on a*

$$\text{Var}(\hat{\theta}_n^A) \leq \text{Var}(\hat{\theta}_{2n}).$$

Pour obtenir la même variance avec l'estimateur classique il faudrait doubler l'échantillon. Au final, on évalue autant de fois la fonction f avec les deux estimateurs. Mais il est moins coûteux de calculer $a - X$ que de simuler une nouvelle fois X , avec une variance qui sera généralement strictement inférieure.

Exemple 4. *Les deux exemples les plus communs de variables aléatoires antithétiques sont :*

- Si $X \sim \mathcal{N}(0, 1)$, $X \stackrel{\mathcal{L}}{=} -X$;
- Si $U \sim \mathcal{U}([0, 1])$, $U \stackrel{\mathcal{L}}{=} 1 - U$.

Exercice 5. *On estime à nouveau via Monte Carlo les quantités de l'exercice 2,*

- $\mathbb{E}[\sin(U)]$ où $U \sim \mathcal{U}([0, \pi])$,
- $\mathbb{E}[\log(X)]$ où $X \sim \mathcal{E}(1)$,
- $\mathbb{E}[(e^{-\frac{\sigma^2}{2} + \sigma Z} - 1)^+]$ où $Z \sim \mathcal{N}(0, 1)$ et $\sigma = 0.2$,

mais en utilisant cette fois des variables aléatoires antithétique. On utilisera $n = 10^6$ et on donnera une estimation de l'écart-type pour chacune des estimations, qu'on comparera à la version précédente sans utilisation de variable antithétique. Pour la simulation de Z , utilisera la fonction `random.normal` du package `numpy` (`rnorm` dans `R`). Comparez les écarts-types et conclure.

3.2 Variable de contrôle

Soit $f(X)$ une variable aléatoire dont on souhaite calculer l'espérance. On suppose qu'on connaît l'espérance de $g(X)$ et on pose $m := g(X)$. Soit $a \in \mathbb{R}$. On pose l'estimateur :

$$\hat{\theta}_n^{C,a} = \frac{1}{n} \sum_{k=1}^n f(X_k) - a(g(X_k) - m).$$

Cet estimateur demeure sans biais et on a

$$\text{Var}(\hat{\theta}_n^{C,a}) = \frac{\sigma_{f(X)}^2 + a^2 \sigma_{g(X)}^2 - 2a \text{Cov}(f(X), g(X))}{n}.$$

Lemme 3. La plus petite variance de $\hat{\theta}_n^{C,a}$ est atteinte pour

$$a^* := \frac{\text{Cov}(f(X), g(X))}{\sigma_{g(X)}^2},$$

et dans ce cas

$$\text{Var} \left(\hat{\theta}_n^{C,a^*} \right) = \frac{\sigma_{f(X)}^2 - \frac{\text{Cov}(f(X), g(X))^2}{\sigma_{g(X)}^2}}{n} = \frac{\sigma_{f(X)}^2 (1 - \rho(f(X), g(X))^2)}{n}$$

Démonstration. Nous avons un polynôme du second degré en a , son minimum est atteint lorsque sa dérivée s'anule, c'est à dire en

$$2a^* \sigma_{g(X)}^2 - 2\text{Cov}(f(X), g(X)) = 0,$$

puis on en déduit le résultat. □

Remarque 4. Le cas $a = 0$ revient au cas sans variable de contrôle. On a une amélioration stricte si $a^* \neq 0$ c'est à dire si $\text{Cov}(f(X), g(X)) \neq 0$.

Exercice 6. Soit $X \sim \mathcal{N}(0, 1)$. On souhaite estimer $\mathbb{E}(|X|)$. Donnez une estimation et son écart-type avec 10^6 simulations de Monte Carlo en utilisant la variable de contrôle X^2 . Comparez au cas sans utiliser de variable de contrôle.

Exercice 7. On estime à nouveau via Monte Carlo les quantités de l'exercice 2,

- $\mathbb{E}[\sin(U)]$ où $U \sim \mathcal{U}([0, \pi])$,
- $\mathbb{E}[\log(X)]$ où $X \sim \mathcal{E}(1)$,
- $\mathbb{E}[(e^{-\frac{\sigma^2}{2} + \sigma Z} - 1)^+]$ où $Z \sim \mathcal{N}(0, 1)$ et $\sigma = 0.2$,

mais en introduisant des variables de contrôle simples dont vous connaissez l'espérance (un polynôme au plus d'ordre 2 pour les deux premiers, et on rappelle que $\mathbb{E}(e^{sZ}) = e^{s^2/2}$ pour le troisième).

Pour simuler la loi exponentielle, on pourra utiliser la fonction `random.exponential` du package `numpy` (`rexp` dans `R`).

3.3 Fonction d'importance ou échantillonnage préférentiel

Pour rappel, on souhaite calculer :

$$\theta := \mathbb{E}[f(X)] = \int_{\mathbb{R}^d} f(x) d\mathbb{P}_X(x).$$

Si \mathbb{P}_X admet la densité f_X par rapport à la mesure μ , et si h est une densité par rapport à μ , alors

$$\theta = \mathbb{E}[f(X)] = \int_{\mathbb{R}^d} f(x) f_X(x) d\mu(x) = \int_{\mathbb{R}^d} \left[\frac{f(x) f_X(x)}{h(x)} \right] h(x) d\mu(x) = \mathbb{E}^h \left[\frac{f(X) f_X(X)}{h(X)} \right].$$

Ici, \mathbb{E}^h est l'opérateur d'espérance avec X qui a pour densité h .

On pose l'estimateur

$$\hat{\theta}_n^{I,h} = \frac{1}{n} \sum_{k=1}^n \frac{f(X_k) f_X(X_k)}{h(X_k)},$$

où $(X_k)_{1 \leq k \leq n}$ est une suite i.i.d. de densité h . L'estimateur demeure sans biais et

$$\text{Var} \left(\hat{\theta}_n^{I,h} \right) = \frac{\text{Var}^h \left(\frac{f(X)f_X(X)}{h(X)} \right)}{n}.$$

Le cas le plus utile est lors du calcul d'évènement rare ou d'espérance où la variable aléatoire n'est pas nulle uniquement lors d'un évènement rare. Dans ce cas, on change la loi de telle sorte à ce que la variable aléatoire ne soit pas souvent nulle sous la densité h .

Exercice 8. *On souhaite calculer*

$$\theta := \mathbb{E} \left((e^{-\frac{\sigma^2}{2} + \sigma Z} - K)^+ \right),$$

avec : $\sigma = 0.2, K \in \{1, 2, 3\}$.

— Implémentez sous Python une méthode avec un échantillonnage préférentiel avec h la densité d'une $\mathcal{N}(\mu, 1)$ avec

$$\mu := \frac{\log(K)}{\sigma} + \frac{\sigma}{2}.$$

— La valeur de μ est choisie afin que la partie positive soit nulle si et seulement si $Z \leq \mu$ (et $\mathbb{P}^h(Z \leq \mu) = \frac{1}{2}$). Déduisez-en une amélioration simple de votre estimateur en utilisant une symétrie et comparez.

Remarque 5. *La méthode utilisée dans le dernier point de l'exercice précédent peut être vue comme une utilisation combinée de variables antithétiques ou comme application particulière des méthodes de stratification de la section suivante.*

3.4 Méthode de stratification

Soit $(A_i)_{1 \leq i \leq \kappa}$ une partition de \mathbb{R}^d . Alors

$$\theta := \mathbb{E}(f(X)) = \sum_{i=1}^{\kappa} \mathbb{P}(X \in A_i) \mathbb{E}(f(X) \mid X \in A_i).$$

On suppose connu $p_{A_i} := \mathbb{P}(X \in A_i)$ pour $1 \leq i \leq \kappa$ et on va estimer les espérances conditionnelles à un évènement. Soit $(n_{A_i})_{1 \leq i \leq \kappa}$ et n tels que

$$n = \sum_{i=1}^{\kappa} n_{A_i}.$$

On va allouer exactement n_{A_i} simulations afin d'estimer $\mathbb{E}(f(X) \mid X \in A_i)$ pour $1 \leq i \leq \kappa$. Soit $(X_k^i)_{1 \leq k \leq n_{A_i}}$ des variables aléatoires indépendante de loi $\mathbb{P}_{X \mid X \in A_i}$, pour $1 \leq i \leq \kappa$. On pose l'estimateur :

$$\hat{\theta}_n^S := \sum_{i=1}^{\kappa} \frac{p_{A_i}}{n_{A_i}} \sum_{k=1}^{n_{A_i}} f(X_k^i).$$

L'estimateur demeure sans biais et sa variance est

$$\text{Var} \left(\hat{\theta}_n^S \right) = \sum_{i=1}^{\kappa} \frac{p_{A_i}^2 \sigma_{f(X) \mid A_i}^2}{n_{A_i}}.$$

Lemme 4. Soit n le nombre total de simulations, le choix des $(n_{A_i})_{1 \leq i \leq \kappa}$ qui minimise la variance est

$$n_{A_i} = n \frac{p_{A_i} \sigma_{f(X)|A_i}}{\sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i}}.$$

On a alors

$$\text{Var}(\hat{\theta}_n^S) = \frac{1}{n} \left(\sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i} \right)^2.$$

Démonstration. On souhaite minimiser en $(n_{A_i})_{1 \leq i \leq \kappa}$ la quantité

$$\text{Var}(\hat{\theta}_n^S) = \sum_{i=1}^{\kappa} \frac{p_{A_i}^2 \sigma_{f(X)|A_i}^2}{n_{A_i}}.$$

sous la contrainte $n = n_{A_1} + \dots + n_{A_\kappa}$. On introduit le langragien $\lambda \in \mathbb{R}$, la condition du premier ordre donne, pour tout $1 \leq i \leq \kappa$,

$$-\frac{p_{A_i}^2 \sigma_{f(X)|A_i}^2}{n_{A_i}^2} + \lambda = 0 \iff n_{A_i} = \frac{p_{A_i} \sigma_{f(X)|A_i}}{\sqrt{\lambda}}.$$

Ainsi $n_{A_i} \propto p_{A_i} \sigma_{f(X)|A_i}$, on en déduit le résultat en réinjectant la contrainte $n = n_{A_1} + \dots + n_{A_\kappa}$. La variance devient

$$\text{Var}(\hat{\theta}_n^S) = \frac{1}{n} \sum_{i=1}^{\kappa} p_{A_i}^2 \sigma_{f(X)|A_i}^2 \frac{\sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i}}{p_{A_i} \sigma_{f(X)|A_i}} = \frac{1}{n} \left(\sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i} \right)^2.$$

□

Lemme 5. L'estimateur par stratification optimale $\hat{\theta}_n^S$ est meilleur que l'estimateur $\hat{\theta}_n$

Démonstration. On a

$$\mathbb{E}(f(X)^2) = \mathbb{E} \left(\sum_{i=1}^{\kappa} \mathbf{1}_{\{X \in A_i\}} f(X)^2 \right) = \sum_{i=1}^{\kappa} p_{A_i} \mathbb{E}(f(X)^2 | X \in A_i).$$

On en déduit, en utilisant des inégalités de convexité

$$\begin{aligned} n \text{Var}(\hat{\theta}_n) &= \sum_{i=1}^{\kappa} p_{A_i} \mathbb{E}(f(X)^2 | X \in A_i) - \left(\sum_{i=1}^{\kappa} p_{A_i} \mathbb{E}(f(X) | X \in A_i) \right)^2 \\ &\geq \sum_{i=1}^{\kappa} p_{A_i} \mathbb{E}(f(X)^2 | X \in A_i) - \sum_{i=1}^{\kappa} p_{A_i} \mathbb{E}(f(X) | X \in A_i)^2 \\ &\geq \sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i}^2 \geq \left(\sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i} \right)^2 = n \text{Var}(\hat{\theta}_n^S). \end{aligned}$$

□

Malheureusement, en pratique, on ne peut pas connaître les poids optimaux car on ne connaît pas les $\sigma_{f(X)|A_i}^2$.

Sans aucune information sur les $\sigma_{f(X)|A_i}^2$, on peut faire comme s'ils étaient constants et prendre :

$$n_{A_i} = n p_{A_i}.$$

Dans ce cas l'estimateur reste plus efficace que l'estimateur classique de Monte Carlo.

Lemme 6. Lorsque le nombre de simulations par strate est :

$$n_{A_i} = np_{A_i},$$

On a :

$$\text{Var}(\hat{\theta}_n^S) = \frac{1}{n} \sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i}^2.$$

En particulier :

$$\text{Var}(\hat{\theta}_n^S) \leq \text{Var}(\hat{\theta}_n).$$

Démonstration.

$$\text{Var}(\hat{\theta}_n^S) = \sum_{i=1}^{\kappa} \frac{p_{A_i}^2 \sigma_{f(X)|A_i}^2}{n_{A_i}} = \frac{1}{n} \sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i}^2.$$

Dans la démonstration du lemme 5, un résultat intermédiaire était que

$$\text{Var}(\hat{\theta}_n) \geq \frac{1}{n} \sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i}^2,$$

ce qui permet de conclure. □

Il est possible en pratique de faire quelques simulations dans chaque strate, afin d'estimer les $\sigma_{f(X)|A_i}$ puis d'utiliser ces estimations pour en déduire le nombre de simulations à effectuer en utilisant le Lemme 5.

Une difficulté demeure : pourvoir simuler les lois conditionnelles $X | X \in A_i$.

Exemple 5. Soit U une loi uniforme sur $[0, 1]$. Soit $\kappa \in \mathbb{N}^*$ et

$$A_i := \left[\frac{i-1}{\kappa}, \frac{i}{\kappa} \right], \quad 1 \leq i \leq \kappa.$$

Alors $U | U \in A_i \sim \mathcal{U}(A_i)$. On pose

$$U_i = \frac{i-1}{\kappa} + \frac{U}{\kappa}.$$

On a bien par construction $U_i \sim \mathcal{U}(A_i)$. De plus $p_{A_i} = \frac{1}{\kappa}$.

Dans l'exemple précédent, si on simule X via l'inverse de la fonction de répartition, c'est à dire $X = F_X^{\leftarrow}(U)$, il suffit de stratifier U . Il est également possible de choisir des intervalles pour X plutôt que U .

Exemple 6. Soit F_X une fonction de répartition et on pose $X = F_X^{\leftarrow}(U)$ où $U \sim \mathcal{U}([0, 1])$. Soit $(a_i)_{1 \leq i \leq \kappa}$ une suite de réels (éventuellements égaux à $-\infty$ ou $+\infty$) strictement croissante telle que $\mathbb{P}(X \in [a_0, a_\kappa]) = 1$.

$$A_i := [F(a_{i-1}), F(a_i)[$$

Alors, $U | U \in A_i \sim \mathcal{U}(A_i)$. On pose

$$U_i = F(a_{i-1}) + (F(a_i) - F(a_{i-1}))U.$$

On a bien par construction $U_i \sim \mathcal{U}(A_i)$. De plus $p_{A_i} = F(a_i) - F(a_{i-1})$. On a également :

$$F_X^{-1}(U_i) \stackrel{\mathcal{L}}{=} X | X \in [a_{i-1}, a_i], \quad 1 \leq i \leq \kappa.$$

Enfin, un cas particulier intéressant est lorsqu'on connaît l'espérance conditionnelle sur certains A_i .

Lemme 7. *On suppose qu'on connaît $m_i := \mathbb{E}(f(X) \mid X \in A_i)$ pour $i < \alpha$ avec $\alpha \geq 2$. Dans ce cas, on pose*

$$n_{A_i} = \begin{cases} 0 & 1 \leq i < \alpha, \\ \frac{np_{A_i}}{\sum_{j=\alpha}^{\kappa} p_{A_j}} & \alpha \leq i \leq \kappa. \end{cases}$$

On pose

$$\hat{\theta}_n^S := \sum_{i < \alpha} p_{A_i} m_i + \sum_{i=\alpha}^{\kappa} \frac{p_{A_i}}{n_{A_i}} \sum_{k=1}^{n_{A_i}} f(X_k^i).$$

Alors

$$\text{Var}(\hat{\theta}_n^S) = \sum_{i=\alpha}^{\kappa} \frac{p_{A_i}^2 \sigma_{f(X)|A_i}^2}{n_{A_i}} = \frac{\sum_{i=\alpha}^{\kappa} p_{A_i}}{n} \sum_{i=\alpha}^{\kappa} p_{A_i} \sigma_{f(X)|A_i}^2.$$

Exercice 9. *On souhaite calculer*

$$\theta := \mathbb{E}\left(\left(e^{-\frac{\sigma^2}{2} + \sigma Z} - K\right)^+\right),$$

avec : $\sigma = 0.2, K \in \{1, 2, 3\}$.

Proposez un estimateur avec $\kappa = 2$ strates et $\kappa = 3$ strates (pour ce dernier cas, on posera $a_2 = a_1 + 1$) et implémentez sous Python.

4 Simulation de processus stochastiques

4.1 Simulation d'un mouvement brownien

L'objectif est de simuler un mouvement brownien $(W_t)_{0 \leq t \leq T}$. Toutefois, il ne sera pas possible numériquement d'avoir une trajectoire sur $[0, T]$. L'objectif est de simuler seulement pour des instants strictement croissants $(t_i)_{0 \leq i \leq n}$ dans $[0, T]$ avec $t_0 = 0$ et $t_n = T$:

$$(W_{t_i})_{0 \leq i \leq n}.$$

Rappelons la définition d'un mouvement brownien.

Définition 2 (Mouvement brownien). *Un processus W est un mouvement brownien si, en posant $\mathcal{F}_t := \sigma(W_s, s \leq t)$,*

- $W_0 = 0$ p.s.
- $t \mapsto W_t$ est continue p.s.
- $W_t - W_s \perp\!\!\!\perp \mathcal{F}_s$ pour $t \geq s$,
- $W_t - W_s \sim \mathcal{N}(0, t - s)$ pour $t \geq s$.

Soit $(Z_i)_{0 \leq i \leq n}$ une suite de variables aléatoires indépendantes telle que

$$\begin{aligned} Z_0 &= 0, \\ Z_i &\sim \mathcal{N}(0, t_i - t_{i-1}), \quad 1 \leq i \leq n. \end{aligned}$$

On définit :

$$\begin{aligned} W_0^n &:= Z_0, \\ W_{t_{i+1}}^n &:= W_{t_i}^n + Z_{i+1}, \quad 0 \leq i \leq n-1, \end{aligned}$$

c'est-à-dire

$$W_{t_i}^n = \sum_{j=0}^i Z_j, \quad 0 \leq i \leq n.$$

Proposition 3. Soit $(W_t)_{t \in [0, T]}$ un mouvement brownien. Pour toute suite $(t_i)_{0 \leq i \leq n}$,

$$(W_{t_i})_{0 \leq i \leq n} \stackrel{\mathcal{L}}{=} (W_{t_i}^n)_{0 \leq i \leq n}.$$

En général, on prendra un pas de temps constant $\Delta t := t_i - t_{i-1}$ pour tout $1 \leq i \leq n$.

Remarque 6. Le mouvement brownien est symétrique, c'est à dire qu'on a

$$W \stackrel{\mathcal{L}}{=} -W.$$

Cette propriété pourra être utilisée comme réduction de variance par variable antithétique.

Exercice 10. Soit $(W_t)_{t \geq 0}$ un mouvement brownien.

- Simulez $n = 10^4$ trajectoires avec un pas de temps de $\Delta t = 0,01$ sur $[0, T]$ (on choisira $T = 1$).
- En déduire un estimateur de $\mathbb{E} [\sup_{0 \leq t \leq T} W_t]$ et donnez sa précision (sans le biais).
- Utilisez $\Delta t = 10^{-4}$ et comparez. On utilisera ce pas pour la suite de l'exercice.
- Proposez un estimateur qui s'appuie sur un processus antithétique et comparez la précision.
- Proposez un estimateur qui s'appuie sur une variable de contrôle et comparez la précision : on utilisera W_T et $\int_0^T W_s ds$.

4.2 Simulation d'une équation différentielle stochastique

4.2.1 Simulation exacte de processus d'Itô

Soit le processus $(X_t)_{t \geq 0}$ solution de

$$X_t = x + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s,$$

où $\mu : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ et $\sigma : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions Lipschitz en la seconde variable uniformément par rapport à la première, à croissance au plus linéaire en la seconde variable uniformément par rapport à la première, et W est un mouvement brownien. Soit $\mathcal{F}_t := \sigma(W_s, s \leq t)$. Alors le processus X est \mathcal{F} -adapté et Markovien relativement à \mathcal{F} , c'est à dire

$$\mathcal{L}((X_u)_{u \geq t} \mid \mathcal{F}_t) = \mathcal{L}((X_u)_{u \geq t} \mid X_t)$$

On pose :

$$\begin{aligned} X_0^n &:= x, \\ X_{t_i}^n &:= Z_i, \quad 1 \leq i \leq n, \end{aligned}$$

où Z_i suit la loi $\mathcal{L}(X_{t_i}^n \mid X_{t_{i-1}}^n)$.

Proposition 4. Pour toute suite $(t_i)_{0 \leq i \leq n}$,

$$(X_{t_i})_{0 \leq i \leq n} \stackrel{\mathcal{L}}{=} (X_{t_i}^n)_{0 \leq i \leq n}.$$

Exemple 7 (Processus de Vašíček). Un processus de Vašíček est la solution de l'équation différentielle stochastique

$$X_t = x + \int_0^t -a(X_s - \mu) ds + \int_0^t \sigma dW_s,$$

avec $a, \sigma > 0$ et $\mu \in \mathbb{R}$. Le processus se réécrit :

$$X_t = \mu + (x - \mu)e^{-at} + \sigma e^{-at} \int_0^t e^{as} dW_s.$$

En particulier, on a, en loi, si pour tout i , $\Delta t = t_{i+1} - t_i$,

$$X_{t_{i+1}} = \mu + (X_{t_i} - \mu)e^{-a\Delta t} + \sigma e^{-a\Delta t} \int_0^{\Delta t} e^{as} dW_s.$$

Il vient

$$X_{t_{i+1}} | X_{t_i} \sim \mathcal{N} \left(\mu + (X_{t_i} - \mu)e^{-a\Delta t}, \frac{\sigma^2}{2a} (1 - e^{-2a\Delta t}) \right)$$

Démonstration. □

Exercice 11 (Estimation d'un Zéro Coupon). Soit $(r_t)_{t \geq 0}$ un processus de Vašíček qui représente le taux instantané sous la probabilité risque-neutre \mathbb{Q} (on se place en marché complet en absence d'opportunité d'arbitrage). Soit $P(t, T)$ la valeur d'un zéro coupon de maturité T à la date t . On rappelle que

$$\left(P(t, T) e^{-\int_0^t r_s ds} \right)_{t \geq 0}$$

est une martingale sous \mathbb{Q} . En conséquence,

$$P(t, T) = \mathbb{E}^{\mathbb{Q}} \left(e^{-\int_t^T r_s ds} | \mathcal{F}_t \right).$$

On se place en date 0 et on observe $r_0 = 0.01$. On pose $a = 0.25$, $\mu = 0.025$, $\sigma = 0.0075$, $T = 2$. On souhaite estimer $P(0, T)$ et on s'appuiera sur une simulation exacte de la trajectoire. On utilisera $n = 10^4$ et $\Delta t = T \times 10^{-3}$. On calculera à chaque fois l'erreur sous forme d'un intervalle de confiance à 95% (sans le biais).

- Proposez une méthode de Monte Carlo simple pour estimer $P(0, t)$ sans méthode de réduction de variance.
- Proposez une méthode de réduction de variance par variable antithétique.
- Proposez une méthode de réduction de variance par variable de contrôle.

Remarque : dans le cas simple de cet exercice, il est possible de calculer explicitement $P(0, T)$, voir en Annexes.

4.2.2 Schéma d'Euler

Soit le processus $(X_t)_{t \geq 0}$ solution de

$$X_t = x + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s,$$

où $\mu : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ et $\sigma : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions Lipschitz en la seconde variable uniformément par rapport à la première, à croissance au plus linéaire en la seconde variable uniformément par rapport à la première, et W est un mouvement brownien. On introduit le processus X^n aux temps $(t_i)_{0 \leq i \leq n}$ avec $X_0^n = x$ par :

$$X_{t_{i+1}}^n = X_{t_i}^n + \mu(t_i, X_{t_i}^n)(t_{i+1} - t_i) + \sigma(t_i, X_{t_i}^n)(W_{t_{i+1}} - W_{t_i}).$$

On a donc en loi :

$$X_{t_{i+1}}^n \mid \mathcal{F}_{t_i} \sim \mathcal{N}(X_{t_i}^n + \mu(t_i, X_{t_i}^n)(t_{i+1} - t_i), \sigma(t_i, X_{t_i}^n)^2(t_{i+1} - t_i))$$

Le processus ci-dessus n'est défini qu'aux instants $(t_i)_{0 \leq i \leq n}$. Il est possible de le prolonger par interpolation linéaire sur $]t_i, t_{i+1}[$ par :

$$X_t^n := X_{t_i}^n + \frac{t - t_i}{t_{i+1} - t_i} (X_{t_{i+1}}^n - X_{t_i}^n).$$

Proposition 5. On pose $t_i = \frac{iT}{n}$, c'est-à-dire $\Delta t = \frac{T}{n}$. Pour tout $p \geq 1$, il existe une constante C_p telle que

$$\sup_{0 \leq t \leq T} \mathbb{E} [|X_t^n - X_t|^{2p}] + \mathbb{E} \left[\sup_{0 \leq i \leq n} |X_{t_i}^n - X_{t_i}|^{2p} \right] \leq \frac{C_p}{n^p}$$

De plus, pour tout $\beta < \frac{1}{2}$,

$$\lim_{n \rightarrow +\infty} n^\beta \sup_{0 \leq t \leq T} |X_t^n - X_t| = 0.$$

Démonstration. Admis. □

Exercice 12 (Estimation du prix d'un Call en volatilité locale). Soit $(S_t)_{t \geq 0}$ le processus de prix d'un actif sous la probabilité risque-neutre et caractérisé (admis) par l'équation différentielle stochastique :

$$dS_t = rS_t dt + \sigma S_t^\gamma dW_t,$$

avec $\frac{1}{2} \leq \gamma < 1$ et $S_0 > 0$. On admettra qu'on peut appliquer le schéma d'Euler. Le prix d'un Call de maturité $T > 0$ et de prix d'exercice $K > 0$ est

$$C(T) = e^{-rT} \mathbb{E} [(S_T - K)^+]$$

On pose $r = 0$, $\sigma = 0.2$, $\gamma = 0.8$, $S_0 = 1$, $K = 1.2$ et $T = 1$.

- Proposez une méthode de Monte Carlo simple pour estimer $C(T)$ sans méthode de réduction de variance.
- Proposez une méthode de réduction de variance par variable antithétique.
- Proposez une méthode de réduction de variance par variable de contrôle s'appuyant sur le modèle de Black-Scholes.

4.2.3 Schéma de Milstein

Soit le processus $(X_t)_{t \geq 0}$ solution de

$$X_t = x + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s,$$

où $\mu : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ et $\sigma : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions Lipschitz en la seconde variable uniformément par rapport à la première, à croissance au plus linéaire en la seconde variable uniformément par rapport à la première, et W est un mouvement brownien. On introduit le processus X^n aux temps $(t_i)_{0 \leq i \leq n}$ avec $X_0^n = x$ par :

$$\begin{aligned} X_{t_{i+1}}^n &= X_{t_i}^n + \mu(t_i, X_{t_i}^n)(t_{i+1} - t_i) + \sigma(t_i, X_{t_i}^n)(W_{t_{i+1}} - W_{t_i}) \\ &\quad + \frac{1}{2} \sigma(t_i, X_{t_i}^n) \partial_x \sigma(t_i, X_{t_i}^n) ((W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1} - t_i)). \end{aligned}$$

Le processus ci-dessus n'est défini qu'aux instants $(t_i)_{0 \leq i \leq n}$. Il est possible de le prolonger par interpolation linéaire sur $]t_i, t_{i+1}[$ par :

$$X_t^n := X_{t_i}^n + \frac{t - t_i}{t_{i+1} - t_i} (X_{t_{i+1}}^n - X_{t_i}^n).$$

Proposition 6. On pose $t_i = \frac{iT}{n}$. Pour tout $p \geq 1$, il existe une constante C_p telle que

$$\mathbb{E} \left[\sup_{0 \leq i \leq n} |X_{t_i}^n - X_{t_i}|^p \right] \leq \frac{C_p}{n^p}$$

De plus, pour tout $\beta < 1$,

$$\lim_{n \rightarrow +\infty} n^\beta \sup_{0 \leq t \leq T} |X_t^n - X_t| = 0.$$

Démonstration. Admis. □

Exercice 13 (Estimation du prix d'un Call en volatilité locale). Refaites l'exercice 12 en utilisant un schéma de Milstein et comparez.

4.3 Application à l'évaluation de produits dérivés

Exercice 14. Soit $(W_t)_{0 \leq t \leq T}$ un \mathbb{Q} - mouvement brownien sur $[0, T]$ et

$$S_t = S_0 e^{(r - \frac{\sigma^2}{2})t + \sigma W_t}.$$

On pose

$$C_T^a = \left(\frac{1}{T} \int_0^T S_t dt - K \right)^+.$$

On prendra $S_0 = 1, r = 0.01, \sigma = 0.2$ et $T = 1$. On fixera $n = 10^5$ et $\Delta t = 0.01$.

— Proposez une méthode de Monte Carlo afin d'estimer

$$C_0^a = e^{-rT} \mathbb{E}^{\mathbb{Q}} (C_T^a).$$

— On introduit la variable :

$$X := \left(\exp \left(\frac{1}{T} \int_0^T \log(S_t) dt \right) - K \right)^+.$$

Proposez une méthode d'estimation de C_0^a avec réduction de variance.

5 Méthodes des différences finies

5.1 Formule de Keynman-Kac

Soit le processus $(X^{t,x})_{\geq t}$ solution de

$$X^{t,x} = x + \int_t^\cdot \mu(s, X_s^{t,x}) ds + \int_t^\cdot \sigma(s, X_s^{t,x}) dW_s,$$

où $\mu : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ et $\sigma : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions Lipschitz et W est un mouvement brownien.

Proposition 7. *Soit $g : \mathbb{R} \rightarrow \mathbb{R}$ une fonction mesurable à croissance polynomiale. On pose :*

$$v(t, x) = \mathbb{E} [g(X_T^{t,x})].$$

Alors, si v est de classe $C^{1,2}$, elle est l'unique solution de l'équation :

$$\begin{cases} \partial_t v(t, x) + \mu(t, x) \partial_x v(t, x) + \frac{1}{2} \sigma(t, x)^2 \partial_{xx}^2 v(t, x) = 0 & (t, x) \in [0, T[\times \mathbb{R} \\ v(T, x) = g(x) & x \in \mathbb{R} \end{cases}$$

Démonstration. Il suffit d'appliquer Itô entre t et $t + h$, de diviser par h et de passer à la limite. Il faudra introduire un temps d'arrêt permettant de localiser le problème pour l'intégrale stochastique. Cela permet de montrer que v satisfait l'équation aux dérivées partielles. L'unicité s'obtient par un théorème de vérification. \square

On a également le résultat plus général.

Proposition 8. *Soient $f : [0, T] \rightarrow \mathbb{R}$ et $r : [0, T] \rightarrow \mathbb{R}$. On définit :*

$$v(t, x) = \mathbb{E} \left[\int_t^T e^{-\int_t^s r(u, X_u^{t,x}) du} f(s, X_s^{t,x}) ds + e^{-\int_t^T r(u, X_u^{t,x}) du} g(X_T^{t,x}) \right].$$

Alors v est l'unique solution de

$$\begin{cases} \partial_t v(t, x) + \mu(t, x) \partial_x v(t, x) + \frac{1}{2} \sigma(t, x)^2 \partial_{xx}^2 v(t, x) - r(t, x) v(t, x) + f(t, x) = 0 & (t, x) \in [0, T[\times \mathbb{R} \\ v(T, x) = g(x) & x \in \mathbb{R} \end{cases}$$

Démonstration. Admis. \square

5.2 Schéma explicite

En réutilisant les fonctions introduites dans la section précédente, l'objectif est de résoudre une équation aux dérivées partielles de la forme

$$\begin{cases} \partial_t v(t, x) + \mu(t, x) \partial_x v(t, x) + \frac{1}{2} \sigma(t, x)^2 \partial_{xx}^2 v(t, x) - r(t, x) v(t, x) + f(t, x) = 0 & (t, x) \in [0, T[\times \mathbb{R} \\ v(T, x) = g(x) & x \in \mathbb{R} \end{cases}$$

On supposera que les fonctions μ et σ vérifient les hypothèses habituelles et que f admet au plus une croissance quadratique en x uniformément en t .

Nous avons besoin d'une maille discrète, sur laquelle nous allons approximer la fonction solution de l'équation aux dérivées partielles. On introduit Δt le pas de temps de discrétisation de T tel que $T/\Delta t \in \mathbb{N}^*$. On introduit le maillage du temps :

$$\mathbf{T}^{\Delta t} := \{t_j := j\Delta t, j \leq T/\Delta t\}$$

Pour $x \in \mathbb{R}$, nous sommes obligés de résoudre sur un espace borné. Soit Δx le pas de temps et $c_-^{\Delta x} \leq c_+^{\Delta x}$ les bornes où nous allons résoudre numériquement l'équation aux dérivées partielles. On introduit le maillage de l'espace :

$$\mathbf{X}^{\Delta x} := \{x_j := j\Delta x, c_-^{\Delta x}/\Delta x \leq j \leq c_+^{\Delta x}/\Delta x\}.$$

On introduit les différences finies pour une fonction φ :

$$\begin{aligned}\partial_{\Delta t}\varphi(t, x) &:= \frac{\varphi(t + \Delta t, x) - \varphi(t, x)}{\Delta t} \\ \partial_{\Delta x}\varphi(t, x) &:= \frac{\varphi(t, x + \Delta x) - \varphi(t, x)}{\Delta x} \\ \partial_{\Delta x}^2\varphi(t, x) &:= \frac{\varphi(t, x + \Delta x) - 2\varphi(t, x) + \varphi(t, x - \Delta x)}{\Delta x^2}.\end{aligned}$$

On évalue la condition terminale en posant $\varphi(T, x) = g(x)$, on pose des conditions aux bords aux points $c^{\Delta x}$ assez éloignée (on pourra y poser $\varphi(t, x) = g(x)$ également ou une autre approximation) car les dérivées impliquent d'aller au-delà en espace. Le schéma explicite consiste à résoudre de manière rétrograde en t :

$$\partial_{\Delta t}\varphi + \mu(t, x)\partial_{\Delta x}\varphi(t + \Delta t, x) + \frac{1}{2}\sigma(t, x)^2\partial_{\Delta x}^2\varphi(t + \Delta t, x) - r(t, x)\varphi(t + \Delta t, x) + f(t, x) = 0$$

La différence finie $\partial_{\Delta t}\varphi$ contient la seule valeur de φ évaluée en t , les autres sont en $t + \Delta t$ ce qui permet de d'obtenir facilement numériquement v connaissant $v(t + \Delta, \cdot)$.

Définition 3 (Schéma explicite par différences finies).

- On pose $\varphi(T, x) = g(x)$ pour $x \in \mathbf{X}^{\Delta x}$,
- Sachant $\varphi(t + \Delta t, \cdot)$ pour un $t \in \mathbf{T}^{\Delta t}$, on déduit de l'équation des différences finies :

$$\begin{aligned}\varphi(t, x) &= \left(1 - \sigma^2(t, x)\frac{\Delta t}{\Delta x^2} - \mu(t, x)\frac{\Delta t}{\Delta x} - r(t, x)\Delta t\right)\varphi(t + \Delta t, x) \\ &+ \left(\frac{1}{2}\sigma^2(t, x)\frac{\Delta t}{\Delta x^2} + \mu(t, x)\frac{\Delta t}{\Delta x}\right)\varphi(t + \Delta t, x + \Delta x) \\ &+ \left(\frac{1}{2}\sigma^2(t, x)\frac{\Delta t}{\Delta x^2}\right)\varphi(t + \Delta t, x - \Delta x) \\ &+ f(t, x)\Delta t.\end{aligned}$$

Proposition 9. Si $(\Delta t, \Delta x) \rightarrow (0, 0)$ et tel qu'on ait toujours (condition de monotonie)

$$1 - \sigma^2(t, x)\frac{\Delta t}{\Delta x^2} - \mu(t, x)\frac{\Delta t}{\Delta x} - r(t, x)\Delta t \geq 0,$$

le schéma est convergent et φ converge simplement vers v .

Démonstration. Admis. □

Concrètement, si $(\Delta t, \Delta x) \rightarrow (0, 0)$ et que $\frac{\Delta t}{\Delta x^2} \rightarrow 0$, la condition de monotonie est satisfaite et le schéma est convergent. La difficulté est de s'assurer que $\frac{\Delta t}{\Delta x^2}$ soit suffisamment petit.

La monotonie peut être testée très facilement sur l'ensemble du schéma.

Remarque 7. On peut écrire le schéma sous la forme matricielle :

$$\varphi(t, x) = (I + A(t))\varphi(t + \Delta t, x) + f(t, x)\Delta t,$$

avec

$$A(t) := \begin{pmatrix} d(t, x_1) & d^+(t, x_1) & 0 & \cdots & 0 & 0 & 0 \\ d^-(t, x_2) & d(t, x_2) & d^+(t, x_2) & \ddots & 0 & 0 & 0 \\ 0 & d^-(t, x_3) & d(t, x_3) & \ddots & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & d(t, x_{n-2}) & d^+(t, x_{n-2}) & 0 \\ 0 & 0 & 0 & \ddots & d^-(t, x_{n-1}) & d(t, x_{n-1}) & d^+(t, x_{n-1}) \\ 0 & 0 & 0 & \cdots & 0 & d^-(t, x_n) & d(t, x_n) \end{pmatrix},$$

où

$$\begin{aligned} d(t, x) &:= -\sigma^2(t, x) \frac{\Delta t}{\Delta x^2} - \mu(t, x) \frac{\Delta t}{\Delta x} - r(t, x)\Delta t, \\ d^+(t, x) &:= \frac{1}{2}\sigma^2(t, x) \frac{\Delta t}{\Delta x^2} + \mu(t, x) \frac{\Delta t}{\Delta x}, \\ d^-(t, x) &:= \frac{1}{2}\sigma^2(t, x) \frac{\Delta t}{\Delta x^2}. \end{aligned}$$

Exercice 15 (Estimation du prix d'un Call en volatilité locale). Refaites l'exercice 12 en utilisant le schéma explicite. On utilisera :

- $\Delta t = 0.00001$,
- $(c_-^{\Delta x}, c_+^{\Delta x}) = (0.5, 2)$,
- $\Delta x = 0.001$.

Testez la condition de monotonie. Observez le comportement du schéma avec $\Delta x = 0.0005$.

5.3 Schéma implicite

Le schéma implicite consiste à prendre les différences finies et l'évaluation de la fonction en espace en t au lieu d'en $t + \Delta t$, c'est à dire de résoudre l'équation :

$$\partial_{\Delta t}\varphi + \mu(t, x)\partial_{\Delta x}\varphi(t, x) + \frac{1}{2}\sigma(t, x)^2\partial_{\Delta x}^2\varphi(t, x) - r(t, x)\varphi(t, x) + f(t, x) = 0$$

On souhaite également résoudre le schéma de manière rétrograde, partant de $\varphi(T, x) = g(x)$. Toutefois, on ne peut pas en déduire immédiatement la valeur de $\varphi(t, x)$, mais on obtient un système linéaire des valeurs de $\varphi(t, \cdot)$ en fonction de celles de $\varphi(t + \Delta t, \cdot)$.

$$\begin{aligned} &\left(1 + \sigma^2(t, x) \frac{\Delta t}{\Delta x^2} + \mu(t, x) \frac{\Delta t}{\Delta x} + r(t, x)\Delta t\right) \varphi(t, x) \\ &+ \left(-\frac{1}{2}\sigma^2(t, x) \frac{\Delta t}{\Delta x^2} - \mu(t, x) \frac{\Delta t}{\Delta x}\right) \varphi(t, x + \Delta x) \\ &+ \left(-\frac{1}{2}\sigma^2(t, x) \frac{\Delta t}{\Delta x^2}\right) \varphi(t, x - \Delta x) \\ &= \varphi(t + \Delta t, x) + f(t, x)\Delta t \end{aligned}$$

L'équation ci-dessous conduit à la représentation matricielle :

$$(I - A(t))\varphi(t, x) = \varphi(t + \Delta t, x) + f(t, x)\Delta t.$$

C'est-à-dire :

$$\varphi(t, x) = (I - A(t))^{-1} (\varphi(t + \Delta t, x) + f(t, x)\Delta t),$$

avec $A(t)$ la matrice définie dans la Remarque 7.

Proposition 10. *Si $(\Delta t, \Delta x) \rightarrow (0, 0)$, le schéma est convergent et φ converge simplement vers v .*

5.4 θ -schéma

Le θ -schéma consiste à prendre une combinaison convexe du schéma explicite et implicite. C'est à dire de résoudre l'équation :

$$\begin{aligned} & \partial_{\Delta t}\varphi + \theta \left(\mu(t, x)\partial_{\Delta x}\varphi(t + \Delta t, x) + \frac{1}{2}\sigma(t, x)^2\partial_{\Delta x}^2\varphi(t + \Delta t, x) - r(t, x)\varphi(t + \Delta t, x) \right) \\ & + (1 - \theta) \left(\mu(t, x)\partial_{\Delta x}\varphi(t, x) + \frac{1}{2}\sigma(t, x)^2\partial_{\Delta x}^2\varphi(t, x) - r(t, x)\varphi(t, x) \right) + f(t, x) = 0 \end{aligned}$$

$$(I - (1 - \theta)A(t))\varphi(t, x) = (I + \theta A(t))\varphi(t + \Delta t, x) + f(t, x)\Delta t,$$

$$\varphi(t, x) = (I - (1 - \theta)A(t))^{-1}(I + \theta A(t))\varphi(t + \Delta t, x) + (I - (1 - \theta)A(t))^{-1}f(t, x)\Delta t,$$

A Simulations - Méthode du rejet

Lemme 8. Soit g une densité de probabilité et $c > 0$ une constante. Alors l'algorithme :

- Tirer X suivant la densité g ,
- Tirer U suivant la loi uniforme sur $[0, cg(X)]$,

permet de tirer (X, U) suivant la loi uniforme sur l'ensemble

$$\mathcal{A} = \{(x, u) \in \mathbb{R}^2, \text{ t.q. } 0 \leq u \leq cg(x)\}.$$

Démonstration. En notant λ_2 la mesure de Lebesgue sur \mathbb{R}^2 , la loi uniforme sur \mathcal{A} a la densité :

$$\frac{1}{\lambda_2(\mathcal{A})} \mathbf{1}_{\mathcal{A}}.$$

On remarque que :

$$\lambda_2(\mathcal{A}) = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{\{0 \leq u \leq cg(x)\}}(x, u) du dx = \int_{\mathbb{R}} cg(x) dx = c.$$

D'autre part, en notant $f_{(X,U)}$ la densité sur \mathbb{R}^2 du couple (X, U) et $f_{U|\{X=x\}}$ la densité conditionnelle de $U | \{X = x\}$, on a (là où $g(x) > 0$, sinon le résultat est trivial) :

$$f_{(X,U)}(x, u) = f_X(x) f_{U|\{X=x\}}(u) = g(x) \frac{1}{cg(x)} \mathbf{1}_{\{0 \leq u \leq cg(x)\}} = c^{-1} \mathbf{1}_{\mathcal{A}}.$$

□

Lemme 9. Soient f et g deux densités de probabilité et $c \geq 1$ une constante telle que $f \leq cg$. Soit $(X, U) \sim \mathcal{U}(\mathcal{A})$. On introduit

$$\mathcal{B} = \{(x, u) \in \mathbb{R}^2, \text{ t.q. } 0 \leq u \leq f(x)\}.$$

Alors

$$(X, U) | U \leq f(X) \sim \mathcal{U}(\mathcal{B}).$$

Démonstration. Dans un premier temps :

$$\begin{aligned} \mathbb{P}(U \leq f(X)) &= c^{-1} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{\{u \leq f(x)\}} \mathbf{1}_{\{0 \leq u \leq cg(x)\}} du dx \\ &= c^{-1} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{\{0 \leq u \leq f(x)\}} du dx = c^{-1} \int_{\mathbb{R}} f(x) dx = c^{-1}. \end{aligned}$$

On en déduit, pour tout borélien $A \in \mathbb{B}(\mathbb{R}^2)$,

$$\begin{aligned} \mathbb{P}[(X, U) \in A | U \leq f(X)] &= \frac{\mathbb{P}[(X, U) \in A, U \leq f(X)]}{\mathbb{P}[U \leq f(X)]} \\ &= \int_A \mathbf{1}_{\{u \leq f(x)\}} \mathbf{1}_{\{0 \leq u \leq cg(x)\}} du dx = \int_A \mathbf{1}_{\{0 \leq u \leq f(x)\}} du dx. \end{aligned}$$

□

Proposition 11 (Méthode du rejet). Soient f et g deux densités de probabilité et $c \geq 1$ une constante telle que $f \leq cg$. La variable aléatoire issue de l'algorithme du rejet :

- (1) Tirer X suivant la densité g ,
- (2) Tirer U suivant la loi uniforme sur $[0, cg(X)]$,
- (3) Si $U \leq f(X)$, conserver X , sinon recommencer à l'étape (1).

a pour densité la fonction f .

Démonstration. On remarque que par le lemme 8, à l'issue des étapes (1) et (2), on a par construction $(X, U) \sim \mathcal{U}(\mathcal{A})$. L'étape (3) n'accepte X que si le couple (X, U) est dans \mathcal{B} , sinon on recommence, ce qui revient à simuler uniformément dans \mathcal{B} par le lemme 9. En vertu du lemme 8, X a alors la loi de probabilité de densité f (cas particulier $c = 1$). \square

Proposition 12. *Le nombre N nécessaire de simulations de couples (X, U) afin de conserver une simulation suit une loi géométrique qui démarre à 1 et de paramètre c^{-1} , c'est-à-dire :*

$$\mathbb{P}(N = n) = c^{-1}(1 - c^{-1})^{n-1}, \quad n \geq 1.$$

En particulier,

$$\mathbb{E}(N) = c.$$

Démonstration. Soit $(X_n, U_n)_{n \geq 1}$ une suite i.i.d. de v.a.r. de loi $\mathcal{U}(\mathcal{A})$, par construction :

$$N = \min\{n \geq 1 \mid (U_n, X_n) \in \mathcal{B}\}.$$

On a par construction :

$$\{N = n\} = \{(U_1, X_1) \notin \mathcal{B}, \dots, (U_{n-1}, X_{n-1}) \notin \mathcal{B}, (U_n, X_n) \in \mathcal{B}\}.$$

Par indépendance,

$$\begin{aligned} \mathbb{P}(N = n) &= \mathbb{P}((U_1, X_1) \notin \mathcal{B}, \dots, (U_{n-1}, X_{n-1}) \notin \mathcal{B}, (U_n, X_n) \in \mathcal{B}) \\ &= \left(\prod_{i=1}^{n-1} \mathbb{P}((U_i, X_i) \notin \mathcal{B}) \right) \mathbb{P}((U_n, X_n) \in \mathcal{B}) \\ &= c^{-1}(1 - c^{-1})^{n-1}. \end{aligned}$$

On a utilisé ci-dessus le fait que, pour tout $i \geq 1$, $\mathbb{P}((U_i, X_i) \in \mathcal{B}) = \mathbb{P}(U_i \leq f(X_i)) = c^{-1}$ qui a été établi dans la démonstration du lemme 9. \square

Exercice 16.

— Montrez que, pour $x \geq 0$,

$$\frac{\sqrt{2}}{\sqrt{\pi}} e^{-\frac{1}{2}x^2} \leq ce^{-x},$$

avec $c \geq \frac{e^{\frac{1}{2}}\sqrt{2}}{\sqrt{\pi}}$,

— En déduire une méthode du rejet pour la simulation de $|X|$ où $X \sim \mathcal{N}(0, 1)$ et l'implémenter sous Python.

— Montrez que le test de rejet peut s'écrire $\log(U) \leq -\frac{1}{2}(X - 1)^2$ avec $U \sim \mathcal{U}([0, 1])$, implémentez ce test et comparez le temps de calcul.

On utilisera `random.exponential` du package `numpy` pour simuler des lois exponentielles (`rexp` en R). Pour comparer le temps de calcul on utilisera la fonction `time` du package `time` (`proc.time` en R) qui donne le temps écoulé en secondes depuis 1970.

B Formule de Black Scholes

$$d_1 := \frac{1}{\sigma\sqrt{T}} [\log(S_0/K) + (r + \sigma^2/2) T]$$

$$d_2 := d_1 - \sigma\sqrt{T}$$

$$C = S_0\Phi(d_1) - Ke^{-rt}\Phi(d_2)$$

$$P = -S_0\Phi(-d_1) + Ke^{-rt}\Phi(-d_2)$$

C Exercice 11 - Prix des Zéros Coupon dans le modèle de Vašíček

La valeur exacte de $P(t, T)$ dans le modèle peut se calculer et est donnée par

$$P(t, T) = e^{A(t, T) - B(t, T)r_t},$$

avec

$$A(t, T) := \left(\mu - \frac{\sigma^2}{2a^2} \right) (B(t, T) - (T - t)) - \frac{\sigma^2}{4a} B(t, T)^2,$$

$$B(t, T) := \frac{1 - e^{-a(T-t)}}{a}.$$